

Model Development Phase Template

Date	10 July 2024
Team ID	SWTID1720001058
Project Title	Panic Disorder Detection
Maximum Marks	4 Marks

Initial Model Training Code, Model Validation and Evaluation Report

The initial model training code will be showcased in the future through a screenshot. The model validation and evaluation report will include classification reports, accuracy, and confusion matrices for multiple models, presented through respective screenshots.

Initial Model Training Code:

jupyter ABCD Last Checkpoint: Last Sunday at 11:00 PM (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel)

```
In [1]: from sklearn.preprocessing import StandardScaler
from sklearn.preprocessing import OneHotEncoder
from sklearn.model_selection import train_test_split
from sklearn.ensemble import GradientBoostingClassifier
from sklearn.metrics import accuracy_score
from sklearn.metrics import confusion_matrix
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
import pandas as pd

In [2]: data = pd.read_csv('panic_disorder_dataset_testing.csv')

In [3]: data
```

```
In [2]: data = pd.read_csv('panic_disorder_dataset_testing.csv')

In [3]: data
```

Out[3]:

	Participant ID	Age	Gender	Family History	Personal History	Current Stressors	Symptoms	Severity	Impact on Life	Demographics	Medical History	Psychiatric History	Substance Use	Coping Mechanisms
0	1	41	Male	Yes	No	High	Shortness of breath	Mild	Mild	Urban	Diabetes	Bipolar disorder	Alcohol	Seeking therapy
1	2	20	Female	Yes	No	Low	Shortness of breath	Mild	Significant	Urban	Asthma	Anxiety disorder	Drugs	Exercise
2	3	32	Male	Yes	Yes	High	Panic attacks	Severe	Mild	Rural	Heart disease	Bipolar disorder	Drugs	Meditation
3	4	41	Female	Yes	Yes	Moderate	Shortness of breath	Moderate	Significant	Urban	Heart disease	Anxiety disorder	NaN	Exercise
4	5	36	Female	Yes	No	High	Chest pain	Severe	Significant	Rural	Asthma	Depressive disorder	NaN	Seeking therapy
...
19995	19996	31	Female	Yes	Yes	High	Chest pain	Moderate	Moderate	Rural	Heart disease	Bipolar disorder	NaN	Exercise
19996	19997	27	Male	No	Yes	High	Panic attacks	Mild	Moderate	Rural	Asthma	Bipolar disorder	Alcohol	Exercise
19997	19998	21	Female	No	No	Low	Dizziness	Moderate	Significant	Urban	Asthma	Anxiety disorder	Alcohol	Socializing
19998	19999	28	Male	No	Yes	Moderate	Dizziness	Mild	Significant	Rural	Heart disease	Anxiety disorder	NaN	Meditation

```
In [6]: data.isnull().sum()

Out[6]: Participant ID      0
Age      0
Gender    0
Family History    0
Personal History  0
Current Stressors 0
Symptoms  0
Severity  0
Impact on Life    0
Demographics      0
Medical History    5001
Psychiatric History 4989
Substance Use      6617
Coping Mechanisms  0
Social Support     0
Lifestyle Factors  0
Panic Disorder Diagnosis 0
dtype: int64

In [7]: data['Symptoms'].value_counts()

Out[7]: Symptoms
Chest pain      4087
Dizziness       4035
Panic attacks   3980
Shortness of breath 3970
Fear of losing control 3928
```

Model Validation and Evaluation Report:

Model	Classification Report	Accuracy	Confusion Matrix
AdaBoost	<pre>print(classification_report(y_test,pred2)) precision recall f1-score support 0 0.96 0.99 0.97 19181 1 0.13 0.04 0.06 819 accuracy 0.55 0.51 0.52 20000 macro avg 0.55 0.51 0.52 20000 weighted avg 0.93 0.95 0.94 20000</pre>	<pre>print("accuracy of Adaboost model:", accuracy_score(y_test, pred2)) accuracy of Adaboost model: 0.9495</pre>	<pre>confusion_matrix(y_test,pred3) array([[19180, 1], [819, 0]], dtype=int64)</pre>
Random Forest	<pre>print(classification_report(y_test,pred)) precision recall f1-score support 0 0.96 1.00 0.98 19181 1 0.33 0.01 0.01 819 accuracy 0.65 0.50 0.58 20000 macro avg 0.65 0.50 0.58 20000 weighted avg 0.93 0.96 0.94 20000</pre>	<pre>str(accuracy_score(y_test,pred)* 100)+ ' %' + ' percentage' '95.88 % percentage'</pre>	<pre>: confusion_matrix(y_test,pred2) : array([[18960, 221], : [786, 33]], dtype=int64)</pre>
Regression	<pre>print(classification_report(y_test, y_pred)) precision recall f1-score support 0 1.00 1.00 1.00 19159 1 1.00 0.99 0.99 841 accuracy 1.00 1.00 1.00 20000 macro avg 1.00 1.00 1.00 20000 weighted avg 1.00 1.00 1.00 20000</pre>	<pre>str(accuracy_score(test_y,pred_y)* 100) 90.56</pre>	<pre>: confusion_matrix(y_test,pred) : array([[19171, 10], : [814, 5]], dtype=int64)</pre>

Gradient Boost

```
print(classification_report(y_test, pred))
```

	precision	recall	f1-score	support
0	0.96	1.00	0.98	19181
1	0.00	0.00	0.00	819
accuracy			0.96	20000
macro avg	0.48	0.50	0.49	20000
weighted avg	0.92	0.96	0.94	20000

```
print("accuracy of gradientboost model:", accuracy_score(y_test, pred))
```

accuracy of gradientboost model: 0.96

